Detecting Adverse Drug Event Safety Signals from MEDLINE Reports: Challenges in Employing Cross-terminology Mapping of MeSH to MedDRA

Abhivyakti Sawarkar, MD, MMSc[1]; Alfred Sorbello, DO, MPH[1]; Anna Ripple, MLS[2]; Olivier Bodenreider MD, PhD[2]

[1]US Food and Drug Administration, Center for Drug Evaluation and Research, Silver Spring, MD, USA; [2]Lister Hill National Center for Biomedical Communications, U.S. National Library of Medicine, Bethesda, MD, USA

Background: The Center for Drug Evaluation and Research/Office of Translational Sciences Data Mining Team at the US Food and Drug Administration (FDA) has recently developed a web-based prototype analytical software tool to automate detection of adverse drug event (ADE) safety signals from MEDLINE reports through quantitative data mining of Medical Subject Heading (MeSH) indexing terms.[1] The prototype tool greatly enhances the FDA's capacity to integrate outputs from MEDLINE with other data mining streams, such as the FDA Adverse Event Reporting System (FAERS), to support multimodal pharmacovigilance. The tool incorporates a system in which MeSH terms are leveraged to extract ADE associations from MEDLINE citations, and the ADE content of a MEDLINE report is conveyed through a single MeSH descriptor.[2] A key element to facilitating interoperability of the data mining outputs from MEDLINE and FAERS is to standardize the representation of ADE content using the Medical Dictionary for Regulatory Activities (MedDRA). In this study, we assess the feasibility of rendering MeSH descriptors for ADEs to MedDRA leveraging the Unified Medical Language System (UMLS) to map MeSH descriptors to equivalent or finer-grained MedDRA preferred terms (PT).

Methods: We manually reviewed the cross-terminology mapping by assessing the completeness and quality of the one-to-one and one-to-many MeSH descriptor-to-MedDRA PT relationships in the context of an arbitrarily selected use case, ciprofloxacin. More specifically, we assessed how well the MedDRA PTs represented the clinical ADE content expressed by the MeSH descriptors in the source MEDLINE citations.

Results: We identified 219 unique drug-AE associations for ciprofloxacin in which the ADE content of the MEDLINE citations was conveyed by unique MeSH descriptors. There were 84 one-to-one MeSH-to-MedDRA PT mappings and 135 one-to-many MeSH-to-MedDRA PT mappings. Manual assessment of the one-to-one mappings confirmed that the linked terms conveyed similar meaning in translating the clinical ADE content of the source citations between the two terminologies. However, the one-to-many mappings included some MeSH descriptors mapped to a multitude of MedDRA PTs (e.g., 'musculoskeletal diseases' mapped to 276 unique MedDRA PTs) that required manual review to select the PT that most reasonably translated the clinical content between the terminologies. Overall, the major categories of cross-terminology challenges involved MEDLINE citations annotated with (1) general MeSH headings (e.g., 'heart diseases') that appeared broad and less distinct compared to the granular mapped MedDRA PTs and that may conflate multiple related major topics, and (2) pre-coordinated descriptors (e.g., 'Drug Hypersensitivity Syndrome') that are conceptually broad in scope and subsume distinct clinically important conditions (e.g., drug reaction with eosinophilia and systemic symptoms (DRESS syndrome)).

Conclusions: This study illustrates some of the challenges in cross-terminology mapping from MeSH to MedDRA and illustrates the limitations inherent in leveraging single MeSH descriptors to convey highly specific ADE content. While one-to-one MeSH descriptors mapped to single MedDRA PTs provide the best links to translate clinical ADE content between the terminologies, the accuracy of translating clinical content between MeSH and MedDRA in one-to-many mappings may be severely impacted by the granularity mismatch between MeSH and MedDRA. Future work will explore text mining of abstracts and full text citations with machine learning to resolve the translation of ADE content between MeSH and MedDRA and thereby support automation.

References

1. Sorbello A, Ripple A, Tonning J, Munoz M, Hasan R, Ly T, Francis H, Bodenreider O. Harnessing scientific literature reports for pharmacovigilance: prototype software analytical tool development and usability testing. Appl Clin Inform. 2017 Mar 22;8(1):291-305.

2. Winnenburg R, Sorbello A, Ripple A, Harpaz, R, Tonning J, Szarfman A, Francis H, Bodenreider O. Leveraging MEDLINE indexing for pharmacovigilance – inherent limitations and mitigation strategies. J Biomed Inform. 2015 Oct;57:425-35.